Available online at www.sciencedirect.com

ScienceDirect

Procedia Computer Science 55 (2015) 405 – 414

Procedia Computer Science

Information Technology and Quantitative Management (ITQM 2015)

A Performance Measure to Support Decision-Making in Agricultural Research Centers in Brazil

Geraldo da Silva e Souza[a], Eliane Gonçalves Gomes[a]*

[a]Brazilian Agricultural Research Corporation, PqEB Av. W3 Norte final, Brasília, DF, 70770-901, Brazil

Abstract

The assessment of productive efficiency of a public research institution is of fundamental importance for its administration. A better management of available resources may be accomplished if managers have at their disposal meaningful quantitative measurements of the production process. In this paper we use Multivariate Analysis and Data Envelopment Analysis to define a performance measure for the research centers of the Brazilian Agricultural Research Corporation. Multiple production indicators are reduced to three output variables by means of maximum likelihood factor analysis. Performance is determined on the basis of this output vector and a three dimensional input vector defined by cost components. We impose restrictions on the optimization algorithm to guarantee usage of all outputs and inputs in the optimal solutions. Types of research centers are compared by using fractional regression models, quasi-maximum likelihood estimation and bootstrap. The analysis also provides a weighting system to compute a goal achievement index and therefore support managerial decision-making.

Peer-review under responsibility of the Organizing Committee of ITQM 2015

Keywords: Factor Analysis; Data Envelopment Analysis; Quasi-Maximum Likelihood Estimation; Agricultural Research; Decision-making.

1. Introduction

It is of importance to the administrators of research institutions to have at their disposal measures and procedures that make feasible an evaluation of the quantum of productivity as well as the technical efficiency of the production process of their institutions. As pointed out in [21], in times of competition and budget constraints a research institution needs to know by how much it may increase its production, with quality, without absorbing additional resources. The quantitative monitoring of the production process allows for an effective administration of the resources available and the observation of predefined research patterns and goals. In this context we developed, at the Brazilian Agricultural Research Corporation – Embrapa, a performance

* Corresponding author. Tel.: +55-61-3448-4476; fax: +55-61-3448-4884.
E-mail address: eliane.gomes@embrapa.br

Peer-review under responsibility of the Organizing Committee of ITQM 2015
doi:10.1016/j.procs.2015.07.094

model based on the input-output data of its research units. The model serves the purpose to evaluate productivity, quantitatively, at relative and absolute levels. The theoretical framework for this model is the analysis of production frontiers. We make intensive use of Data Envelopment Analysis – DEA – models. See, for instance [2-5].

Embrapa's research system currently comprises 42 research centers, or Decision Making Units (DMUs), in the DEA jargon, classified into three types somewhat according to the nature of their operation. These are product (14 research centers), eco-regional (17) and thematic (11) centers. Five of these units were recently created (2010–2012) and are not included in the evaluation system discussed here, so that our sample consists of 37 DMUs. With the participation of the Board of Directors of Embrapa, as well as the administration of each of its research units, 30 outputs, classified into four categories, and three input indicators were identified as being representative of company production actions.

Any sensible performance model will aggregate the output variables, within categories, into a smaller set to allow practical usage of the information available. Aggregation implies the use of a proper weighting system and assumes that performance variables are measured on a homogenous scale. The usage of value judgments to define weights via multicriteria decision methods, as the AHP or the Macbeth for instance, is commonly found in the literature, but has the drawback to be much administration dependent. The use of weighting systems based on principal component analysis or factor analysis to reduce the dimension of the output space is very appealing statistically, but it is also subjective since it relies on a particular rotation of the factor space. Another disadvantage is its dependence on factor scores potentially involving negative factor loadings. Our suggestion is to define a system of weights based on maximum likelihood factor analysis and relative communalities. These weights are always positive and independent of orthogonal rotations in the factor space.

After transformation of input and output variables to a homogeneous scale and reduction of the output space by means of multivariate methods, we assess performance by means of DEA with multipliers restrictions to insure usage of all inputs and outputs in the optimal solution. The lower bounds on these multipliers are found considering the assurance region proposed by [6].

If one is interested in the effects of contextual variables on the DEA performance measurements, such as type of research center or any other exogenous factors, the analysis may be performed by specific regression-like methods. These approaches are known as two-stage estimation and are discussed in detail in [7-10]. The statistical problems in the two-stage approach relate to the cross-sectional correlations induced by the way DEA measures are computed and the potential endogeneity of a contextual variable. Here our concern is in the performance comparisons of three types of research centers. We follow [8] using a quasi-maximum likelihood (QML) method combined with bootstrap.

2. Inputs and Performance Indicators

The performance indicators (outputs) available for Embrapa are classified into four categories: (a) scientific production, (b) production of technical publications, (c) development of technologies, products, and processes, and (d) diffusion of technologies and image. They are measured by number of cases. These indicators were used, in a different form, in [1].

Scientific production entails, mainly, the publication of articles and book chapters. It is required that each item be specified with a complete bibliographical reference. The category of scientific production includes the items: Scientific articles published in refereed journals and book chapters, including both domestic and foreign publications; Articles and summaries published in proceedings of congresses and technical meetings; and Supervision of academic (graduate) dissertations or thesis.

The technical publications category groups publications produced by research centers that focus primarily on agricultural businesses and agricultural production. The category includes: Technical circulars; Research bulletins: Technical communiqués; Periodicals (serial publications not classified in the previous categories); Technical recommendations (aimed at extensionists and farmers in general); and Ongoing research.

The development of technologies, products and processes category comprises indicators related to the efforts made by a research unit to provide its production to society as a final product. Only new technologies, products and processes are considered here. The category includes: Cultivars; Agricultural and livestock processes and practices; Agricultural and livestock inputs; Agro-industrial processes; Machinery (equipment); Scientific methodologies; Software; and, Monitoring, zoning (agro-ecologic or socioeconomic) and mapping.

The diffusion of technologies and image category includes production actions related to Embrapa's efforts to make its products known to the public and to market its image. The category includes the following indicators: Field days; Organization of congresses and seminars; Seminar presentations; Participation in expositions and fairs; Courses offered; College-level training programs; Fellowship holders (refers to the orientation of students, who are the fellowship holders); Folders; Videos; Demonstration units; and Observation units.

The input side of Embrapa's performance process is composed of three factors: Personnel costs (salaries plus labor duties), Operational costs (expenses resulting from the consumption of materials, travel and services, less income from production projects), and Capital (measured by depreciation).

To reduce scale problems, performance indicators and inputs are normalized by the number of employees of each research center. Performance of a research center will be assessed using its ranks on these variables. This procedure turns the units comparable and is robust relative to outliers.

3. Factor Analysis

To study the appropriateness of each of the four output dimensions defined in Section 2 we make use of maximum likelihood estimation and factor analysis. A one-factor model should fit the data in each instance if the dimensions are well defined. References for our discussion on factor analysis here are [11-12]. A vector variable x of dimension p with mean μ and variance-covariance matrix Ω satisfies the k-factor model if one can write $x - \mu = \Lambda f + u$, where $\Lambda(p \times k)$ is a matrix of constants and $f(k \times 1)$ and $u(p \times 1)$ are random. The components of f are the common factors and of u are the specific factors. One assumes $E(f) = 0,\ Var(f) = I,\ E(u) = 0,\ Cov(u_i, u_j) = 0\ \ i \neq j, Cov(f, u) = 0.$ We impose multinormality for (f, u).

Let $\Phi = diag(\varphi_{11}, \ldots, \varphi_{pp})$ denote the variance-covariance matrix of u. We have $x_i - \mu_i = \sum_{j=1}^{k} \lambda_{ij} f_j + u_i$. The variance σ_i^2 of x_i is given by $\sigma_i^2 = \sum_{j=1}^{k} \lambda_{ij}^2 + \varphi_{ii}$. The component $h_i^2 = \sum_{j=1}^{k} \lambda_{ij}^2$ is called the communality and represents the variance of x_i which is shared with the other variables via the common factors. In particular, $\lambda_{ij} = Cov(x_i, f_j)$ is the extent to which x_i depends on the jth common factor. The k-factor model can be expressed by the condition $\Omega = \Lambda\Lambda' + \Phi$ and it is invariant to location and scale transformations on the components of x.

If the k-factor model holds for x then, for any orthogonal matrix $G(k \times k)$, one has $x - \mu = (\Lambda G)(G' f) + u$ and $\Omega = (\Lambda G)(G'\Lambda') + \Phi$. Thus, rotations in the factor space will not lead to distinct formulations of the model. The k-factor remains valid with new factors $G' f$ and factor loadings ΛG.

Let S denote the sample variance-covariance of x. Maximum likelihood estimation of Λ and Φ are obtained maximizing the log likelihood function $-1/2\, n \log|2\pi\Omega| - 1/2\, \mathrm{tr}\, \Omega^{-1} S$ with respect to Λ and Φ. The k-factor multinormal model allows a goodness of fit test. The hypothesis H_k that the k-factor fits the data can be tested using the likelihood ratio test statistic $-2\ln\lambda = np(\hat{a} - \log g - 1)$, where $\hat{a}$ and $\hat{g}$ are the arithmetic and the geometric means of the eigenvalues of $\hat{\Omega}^{-1} S$. Under the null hypothesis H_k, $-2\ln\lambda$ is χ^2 with $s = 1/2(p-k)^2 - 1/2(p+k)$ degrees of freedom. When k=0 (components of x will be independently distributed), the test statistic may be computed using the formula $-n\log|R|$, where R is the correlation matrix.

[13] showed that the chi-squared approximation is improved if *n* is replaced by $n' = n - 1 - 1/6(2p+5) - 2/3k$.

An index of sampling adequacy of the *k*-factor model commonly used is the Kaiser-Meyer-Olkin (*KMO*) statistic [14]. For the *i*th component of *x* it is defined by $KMO_i = \sum_{j \neq i} r_{ij}^2 \Big/ \sum_{j \neq i} r_{ij}^2 + \sum_{k \neq i} b_{ij}^2$, where $R = (r_{ij})$ and $B = (b_{ij})$ is the partial correlation matrix. The global *KMO* is given by $KMO = \sum_i \sum_{j \neq i} r_{ij}^2 \Big/ \sum_i \sum_{j \neq i} r_{ij}^2 + \sum_i \sum_{k \neq i} b_{ij}^2$. In exploratory factor analysis it is suggested that $KMO < 0.50$ is indicative that *R* is not adequate for factor analysis.

The use of ranks with standard multivariate normal methods instead of the original values of the (possibly non-normal) variables endows the procedure with nonparametric properties [15]. Here the use of ranks as production indicators allows aggregation, since they are unit independent. The score of a research centre *i* in the dimension *d* of the performance evaluation comprising *v* variables with corresponding rank vector $(c_1^{id}, \ldots, c_v^{id})$ and based on the *k*-factor model appropriate for that dimension is defined by $y_i^d = \sum_{j=1}^{v} \theta_{jd} c_j^{id}$, $\theta_{jd} = h_{jd}^2 \Big/ \sum_{\tau=1}^{v} h_{\tau d}^2$. Here h_{jd}^2 is the communality of the *j*th variable. The weighting system is invariant to orthogonal transformations of the factor model.

4. DEA Models

Consider a production process composed of *n* DMUs. Each DMU uses varying quantities of *m* different inputs to produce varying quantities of *s* different outputs. Denote by $Y = (y_1, y_2, ..., y_n)$ the $s \times n$ production (output) matrix of the *n* DMUs. The *r*th column of *Y* is the output vector of DMU *r*. Denote by $X = (x_1, x_2, ..., x_n)$ the $m \times n$ input matrix. The *r*th column of *X* is the input vector of DMU *r*. The measure of technical efficiency of production (performance index), under constant returns to scale (CRS) for DMU $o \in \{1, 2, ..., n\}$, denoted $\mathrm{E}^{\mathrm{CR}}(o)$, is the solution of the problem (1). This is called CCR model [16].

$$\begin{aligned} &\max\ \mathrm{E}^{\mathrm{CR}}(o) = \sum_{j=1}^{s} u_j y_j^{o} \Big/ \sum_{i=1}^{m} v_i x_i^{o} \\ &\text{subject to } \ \text{i)} \sum_{i=1}^{m} v_i x_i^{o} = 1,\ \text{ii)} - \sum_{i=1}^{m} v_i x_i^{k} + \sum_{j=1}^{s} u_j y_j^{k} \le 0 \quad \forall k \ \text{ and iii)}\, u_j, v_i > 0 \quad \forall\, j,i \end{aligned} \tag{1}$$

If we look at the coefficients *u* and *v* as input and output prices, we see that the measure of technical efficiency of production is very close to the notion of productivity (output income/input expenditure). Technical efficiency, in this context, basically, is looking for the price system (u, v) for which DMU *o* achieves the best relative productivity ratio. The dual problem of the linear programming problem (1) has an important economic interpretation. This is equivalent to formulation (2).

$$\begin{aligned} &\min_{\theta,\lambda} \theta \\ &\text{subject to} \\ &\text{i) } Y\lambda \ge y_o,\ \ \text{ii) } X\lambda \le \theta x_o \ \text{ and iii) } \lambda \ge 0,\ \theta \text{ free} \end{aligned} \tag{2}$$

The matrix products $Y\lambda$ and $X\lambda$, with $\lambda \ge 0$, represent linear combinations of the columns of *Y* and *X*, respectively, i.e., a sort of weighted averages of output and input vectors. In this way, for each $\lambda \ge 0$ we can generate a new production relation, a new “pseudo” producer. Trivially, the set of DMUs 1, 2,..., *n* is included among those new producers. Making allowance for these newly defined production relationships, the question that the dual intends to answer is: what proportional reduction of inputs $\theta\, x_o$ it is possible to achieve for DMU *o* and still produce at least output vector y_o? The solution $\theta^*(x_o, y_o)$ is the smallest θ with this property.

We can define the concept of technical efficiency of production in a context of fixed inputs instead of fixed outputs, i.e., in a program of output augmentation. In the output augmentation program the question is: what proportional rate ϕ can be uniformly applied to augment the output vector y_o, without increasing the input vector x_o? The solution ϕ^* is the largest ϕ with this property.

Questions of scale can be dealt with properly by imposing restrictions in the linear programming problem. One obtains the variable returns DEA imposing the additional condition $1'\lambda = 1$ on the weight vector λ. This is called the BCC model [17].

DEA scores may attribute unit efficiency to DMUs that are not Pareto efficient in the sense that the multipliers of the corresponding optimal solutions are nonzero. Pareto efficiency is a desirable property, particularly in performance evaluation systems, since it forces all units to effectively use all outputs and inputs. In our application it is awkward to allow for a unit to be efficient with a zero weight in a particular input or output component. In order to force nonzero multipliers, the idea is to determine a small constant $\varepsilon > 0$ and impose ε as a lower bound to the DEA multipliers, as in (3).

$$\max \sum_{j=1}^{s} u_j y_j^o$$
$$\text{subject to} \tag{3}$$
$$\text{i)} \sum_{i=1}^{m} v_i x_i^o = 1,\ \text{ii)} -\sum_{i=1}^{m} v_i x_i^k + \sum_{j=1}^{s} u_j y_j^k \le 0 \quad \forall k \ \text{ and iii)}\, u_j, v_i \ge \varepsilon > 0 \quad \forall\, j, i$$

As discussed in [4, 18] is tempting to choose $\varepsilon = 10^{-5}$ or smaller. A relatively large value of ε may lead to an unfeasible linear programming problem (LPP). A too small value may lead to a solution very close to standard DEA, with multipliers too small to be sensible. Our approach is to choose the largest ε possible yielding a feasible solution. This quantity may be computed using the upper limit of the assurance interval proposed by [6]. This is determined solving the LPP (4) and computing $\varepsilon^* = \min\{\varepsilon_1^*, ..., \varepsilon_n^*\}$. The assurance interval is [0,ε*]. Our choice for the DEA evaluation is input orientation under CRS assumption.

$$\max \ \varepsilon_o$$
$$\text{subject to} \tag{4}$$
$$\text{i)} \sum_{i=1}^{m} v_i x_i^o = 1,\ \text{ii)} -\sum_{i=1}^{m} v_i x_i^k + \sum_{j=1}^{s} u_j y_j^k \le 0 \quad \forall k \ \text{ and iii)}\, u_j \ge \varepsilon_o,\ \ v_i \ge \varepsilon_o$$

5. Two-stage Statistical Analysis

Care should be exercised in the analysis if one is concerned with the statistical inference related to the effect of contextual variables in DEA performance measurements. Firstly, DEA measures are correlated by the very nature of their computations. Secondly, the potential correlation of a covariate with the efficiency index may invalidate the analysis in a manner similar to what happens with the use of ordinary least squares in the presence of endogenous independent variables. See [9] for more details.

Here we are interested in assessing real differences in performance due to type of units. It is a typical analysis of variance problem and we do not expect endogeneity of the classification. In [8] it is suggested the use fractional regression models. Let an observed DEA response $\hat{\theta}$ be dependent on a vector of covariates *w*. They consider a one- and a two-part model, the models differing in the way the efficient units are treated. In the one-part model it is assumed that $E(\hat{\theta} \mid w) = G(w\delta)$, where G(.) is a probability distribution function. The model is well defined even when θ puts positive probability mass at one. The unknown parameter δ is then estimated by quasi-maximum likelihood (QML), maximizing $\sum_{i=1}^{n}\left(\hat{\theta}_i \log\left(G\left(w_i\delta\right)\right)+(1-\theta_i)\log\left(1-G\left(w_i\delta\right)\right)\right)$.

The two-part model uses the whole sample to estimate the model $\text{Prob}\left(\hat{\theta}_i = 1 \middle| w_i\right) = F\left(w_i'\beta\right)$, where β is an unknown parameter vector, and F is a known probability distribution function. For the second part it is assumed $E(\hat{\theta}_i \mid w_i) = G\left(w_i'\delta\right)$ for the responses in (0,1). Typical choices for F and G in both models are the logistic, probit and the log-log distribution functions. We favor the use of the one part model if the sample is comprised of only a few efficient units. Indeed, this is our choice here.

For the one-part model, [19] show that under the correct specification of the mean function $\sqrt{n}\left(\hat{\delta}-\delta\right)\xrightarrow{d} N(0,V)$. V is estimated using (5). The QML estimator is efficient within the class of estimators containing all linear exponential family-based QML and weighted nonlinear least squares estimators [8].

$$\begin{aligned}&\hat{V}=\left(\hat{A}\right)^{-1}\hat{B}\hat{A},\quad \hat{A}=1/n\sum_{i=1}^{n}\left(\hat{g}_i^2/\hat{G}_i(1-\hat{G}_i)\right)w_i'w_i,\quad \hat{B}=1/n\sum_{i=1}^{n}\left(\hat{u}_i^2\hat{g}_i^2/(\hat{G}_i(1-\hat{G}_i))^2\right)w_i'w_i\\&\hat{G}_i=G(w_i'\hat{\delta}),\quad \hat{g}_i=G'(w_i'\hat{\delta}),\quad \hat{u}_i=\hat{\theta}_i-\hat{G}_i\end{aligned} \tag{5}$$

The validity of the asymptotic distributional properties of the QML estimator is also dependent on a condition not pointed out by [8]. This is the notion of Cesaro summability. See [20]. Cesaro summability is basically a property of independent sequences of random variables and implies uniform strong laws of large numbers. A good reference in this regard is [21]. In [19], for example, it is assumed independence. It is not easy to verify Cesaro summability for correlated sequences, as the DEA scores. For this reason, our choice to derive the distributional properties of $\hat{\delta}$ is the bootstrap. We compute $\hat{\delta}$ by QML and draw repeated samples of centered residuals, with replacement. In each sample, and for each DMU, we add a sampled zero mean residual to $G\left(w_i'\hat{\delta}\right)$. Then a new value of $\hat{\delta}$ is computed by QML. The process is repeated 5,000 times.

6. Empirical Results

6.1. Consistency of dimensions and marginal performance indicators

The empirical measures of the KMO type for V101 (scientific articles), V103 (book chapters), V104 (articles in proceedings), V105 (abstracts in proceedings), and V106 (supervision of graduate thesis) are shown in Table 1. This set of variables seems to be well represented by a one-factor factor model. Table 1 also shows the relative communalities (weights). The likelihood ratio test does not indicate misspecification.

The empirical measures of the KMO type for V201 (technical circulars), V202 (technical communiqués), V203 (research bulletins), and V204 (periodicals are shown in Table 2. This set of variables is also well represented by a one-factor factor model. The other variables in this dimension show low values of KMO and were discarded. The likelihood ratio test does not indicate misspecification. Relative communalities for the final set of variables in this dimension (final weights) are also shown in Table 2.

The dimensions of development of technologies, products and processes and of diffusion of technologies and image are not consistent. There are many DMUs that do not perform specific activities in these categories, and an overall factor model is not appropriate. Pooling all the variables in these two dimensions and following a study of the marginal KMOs we found convenient to reduce the dimensionality to one category considering the variables V301 (field days), V302 (organization of congresses and seminars), V303 (seminar presentations), V308 (fellowship holders), V311 (observation and demonstration units), V408 (scientific methodologies), and V411 (monitoring and zoning). If it is of management concern, the discarded variables should be inspected outside the model proposed here. The marginal KMOs are shown for the new dimension in Table 3. The low value of V411 prevents the consistency of a one-factor model, which is the case otherwise. A two-factor model is necessary to represent this dimension. This does not invalidate our weighting system. Relative communalities are therefore computed under this assumption. We choose to include V411 in the model given its environmental importance. The maximum likelihood ratio test accepts the two-factor model. The same test rejects the one-factor model.

Table 1. KMO measures and final weights for the dimension Scientific Production (SCP)

Variable	KMO	Communality	Weights (%)
V101	0.6638	0.8299161	38.4530
V103	0.6840	0.0422336	1.9568
V104	0.5831	0.1245829	5.7724
V105	0.6743	0.6772608	31.3799
V106	0.8322	0.4842674	22.4379

Table 2. KMO measures and final weights for the dimension Technical Publications (TEP)

Variable	KMO	Communality	Weights (%)
V201	0.5850	0.1551894	8.6052
V202	0.6332	0.2946661	16.3390
V203	0.6110	0.3535586	19.6046
V204	0.5773	0.3958007	21.9469

Table 3. KMO measures and final weights for the dimension Other Performance Indicators (OPI)

Variable	KMO	Communality	Weights (%)
V301	0.6057	0.7260228	10.0027
V302	0.6327	0.5810623	8.0055
V303	0.6763	0.5619325	7.7420
V308	0.6657	0.1032986	1.4232
V311	0.5574	0.2014535	2.7755
V408	0.6058	0.3302008	4.5493
V411	0.4957	0.5188184	7.1480

6.2. Performance assessment

Performance indicators and performance measurement were computed as previously explained: inputs are ranks of the corresponding normalized variables; outputs are the average ranks computed with weights defined by relative communalities; performance scores were obtained from the DEA model with multipliers restrictions.

The performance score has rank correlation 0.929 with the classical DEA-CCR measure. The upper bound of the assurance region is 0.002652. The average optimum relative multipliers (weights) for the inputs are 19.4 %, 24.6 % and 56.0 % for labor, operational, and capital costs, respectively. Labor is therefore the “cheapest” component relatively to the inputs shadow prices for the system. Regarding the output components the figures are 66.9 %, 20.3 % and 12.8 % for scientific production, technical publications, and other activities, respectively. Scientific production dominates the performance index. *Ceteris paribus*, DMUs with good scientific production and low values of capital expenses perform better.

The distributions by type of research centers differ and DMUs classified as of type product have superior performance, followed by the eco-regional and the thematic ones. The median performances for these groups are 0.610, 0.421 and 0.260, respectively.

The formal statistical analysis with the QML estimator is provided in Table 4, where it is shown bootstrap standard error and bias corrected confidence intervals [22]. Table 4 assumes a logistic response. The probit and the log-log specifications do not lead to better models or different conclusions. The model fitted assumes the expected performance be dependent of $\mu_i = d_0 + d_1 d_i^1 + d_2 d_i^2$ defined by the indicator variables d_i^1 and d_i^2 of types product and thematic, respectively. The parameter vector estimate is $\hat{d} = (0.3085,\ 0.04644,\ -1.2722)$, suggesting low performance for the thematic DMUs.

Table 4. QML estimation with bootstrap standard errors and bias corrected confidence intervals (mean response is the logistic)

	Parameter	Standard error (QML)	Standard error (bootstrap)	Bias corrected confidence intervals	
				Lower 95%	Upper 95%
d_0	0.30850	0.39346	0.30401	-0.249446	0.934621
d_1	0.04644	0.47428	0.41489	-.0750329	0.886993
d_2	-1.27220	0.42806	0.59135	-2.610639	-.2778407
$d_1 - d_2$	1.31860	0.31395	0.57659	.3464186	2.587331

Table 5 shows tests of normality of the parameters of concern $\hat{d}$ and $\hat{d}_1 - d_2$ using the Kolmogorov-Smirnov test statistic. There are biases and only d_1 passes the normality test. The differences in responses between thematic and product, eco-regional and thematic, and eco-regional and product research centers are assessed by the statistical significance of d_1-d_2, d_2, and d_1, respectively. Product centers have significantly higher performance than thematic. The thematic DMUs perform lower on average than eco-regional. No other differences were found statistically significant at the 5% level. The estimated expected responses are 0.57652, 0.58782, and 0.27614 for eco-regional, product and thematic, respectively.

Table 5. Kolmogorov-Smirnov normality tests

Parameter	D	Prob > D
d_0	0.016871	<0.01
d_1	0.010920	>0.15
d_2	0.034927	<0.01
$d_1 - d_2$	0.034219	<0.01

6.3. Goal achievement indices

The weighting system derived from Factor Analysis based on relative communalities can be used to define goal achievement scores by output dimension for the following evaluation period. In this context it is assumed the existence of a set of performance goals negotiated between the company managers and the local managers.

For output dimension ν, with untransformed variables $x_i,\ i = 1,\ldots,m_\upsilon$ and weights $w_i,\ i = 1,\ldots,m_\upsilon$, the goal achievement index GAI_ν^d for unit *d* is defined by (6), where x_{id}^m is the goal and x_{id}^o is the actual value observed for variable $x_i,\ i = 1,\ldots,m_\upsilon$.

$$GAI_\nu^d = \sum_{i=1}^{m_\upsilon} w_i \frac{x_{id}^o}{x_{id}^m} \quad (6)$$

An overall goal achievements index may be computed aggregating the marginal goal achievement indices using the nonzero DEA multipliers, normalized to sum 1.

7. Summary and Conclusions

The objective of this article was to define a performance index for each of the 37 research centers of Embrapa based on a set of input variables – costs of labor, operation, and capital – and 30 output indicators. The purpose is to aid the administration to monitor and control production of key output variables for the organization, like the production of scientific articles and technology transfers to the agricultural sector.

All variables were rank transformed to reduce size influence, to allow aggregation and to enable a nonparametric multivariate type analysis. The output indicators were then studied by means of maximum likelihood factor analysis, were reduced to 16 variables, and were classified into three output dimensions. The consistency of each output dimension was assessed using empirical measures of adequacy for the factor model and by means of statistical tests based on the likelihood ratio associated with a multivariate normal factor analysis. Relative communalities, derived by the factor models, were used as weights in the computation of average scores for each dimension.

We then used nonparametric efficiency analysis to compute performance scores using a three input – three output DEA model under CRS. An optimal lower bound was used as a restriction on the DEA multipliers forcing nonzero weights for all inputs and outputs in the optimal solutions leading to the performance scores. A fractional regression model with a logistic response was fitted by quasi-maximum likelihood. Biases corrected bootstrap confidence intervals were used to assess differences in performance between types of units.

We also proposed a goal achievement index to monitor the attainment of strategic goals for the output variables, measured in their original scales.

We see three main contributions of our work. Firstly, the proposed weighting system is invariant by rotations of the factor model and is fully data oriented. Weights, in principle, are supposed to be user defined and should reflect the administration’s perception of the relative importance of each variable. Defining weights is a hard and questionable task and we did not succeed with traditional subjective methods. Secondly, we provide a way to assess the consistency of output indicators via multivariate analysis, when reducing dimensionality is deemed important to asses and control performance. Finally, the use of DEA with optimal multipliers restrictions provides a new nonzero system of weights that can be combined with the dimension reduction to provide an additional way to monitor performance and production via goals achievement.

References

[1] Souza GS, Alves E, Avila AFD. Technical efficiency in agricultural research. *Scientometr* 1999; **46**:141–160.

[2] Charnes A, Cooper WW, Lewin AY, Seiford LM. *Data Envelopment Analysis: Theory, Methodology and Application*. Boston: Kluwer Academic Publishers; 1996.

[3] Coelli TJ, Rao DSP, O’Donnell CJ, Battese GE. *An Introduction to Efficiency and Productivity Analysis*. 2nd ed. New York: Springer; 2007.

[4] Cooper WW, Seiford LM, Tone K. *Data Envelopment Analysis: A Comprehensive Text with Models, Applications, References and DEA-Solver Software*. 2nd ed. New York: Springer; 2007.

[5] Cooper WW, Seiford LM, Zhu J. *Handbook on Data Envelopment Analysis*. 2nd ed. New York: Springer; 2011.

[6] Mehrabian S, Jahanshahloo GR, Alirezaee MR, Amin GR. An assurance interval for the non-Archimedean epsilon in DEA models. *Oper Res* 2000; **48(2)**:344–347.

[7] Banker RD, Natarajan R. Evaluating contextual variables affecting productivity using data envelopment analysis. *Oper Res* 2008; **56**:48–58.

[8] Ramalho EA, Ramalho JJS, Henriques PD. Fractional regression models for second stage DEA efficiency analyses. *J Product Anal* 2010; **34**:239–255.

[9] Simar L, Wilson PW. Estimation and inference in two-stage, semi-parametric models of production processes. *J Econom* 2007; **136(1)**:31–64.
[10] Souza GS, Staub RB. Two stage inference using data envelopment analysis efficiency measurements in univariate production models. *IntTrans Oper Res* 2007; **14**:245–258.
[11] Johnson RA, Wichern DM. *Applied Multivariate Statistical Analysis*. 6th ed. New Jersey: Pearson/Prentice Hall; 2007.
[12] Mardia LV, Keni JT, Bibby JM. *Multivariate analysis*. London: Academic Press; 1979.
[13] Bartlett MS. A note on multiplying factors for various chi-squared approximations. *J R Stat Soc* 1954; **16**:296–298.
[14] Kaiser HF. An index of factor simplicity. *Psychom* 1974; **39**:31–36.
[15] Conover MJ. *Practical Nonparametric Statistics*. 3rd ed. New York :Wiley; 1999.
[16] Charnes A, Cooper WW, Rhodes E. Measuring the efficiency of decision-making units. *Eur J Oper Res* 1978; **2**:429–444.
[17] Banker RD, Charnes A, Cooper WW. Some models for estimating technical scale inefficiencies in Data Envelopment Analysis. *Manag Sci* 1984; **30(9)**:1078–1092.
[18] Ali AI, Seiford LM. The mathematical programming approach to efficiency analysis. In: Fried HO, Lovell CAK, Schmidt SS, editors. *The Measurement of Productive Efficiency: Techniques and Applications*, Oxford: Oxford University Press; 1993, p. 120–159.
[19] Papke LE, Wooldridge JM. Econometric methods for fractional response variables with an application to 401(k) plan participation rates. *J Appl Econom* 1996; **11(6)**:619–632.
[20] Gourieroux C, Monfort A, Trognon A. Pseudo-maximum likelihood methods: theory. *Econom* 1984; **52**:681–700.
[21] Gallant AR. *Nonlinear Statistical Models*. 1st ed. New York: Wiley; 1987.
[22] Efron B, Tibshirani RJ. *An Introduction to the Bootstrap*. New York: Chapman & Hall; 1994.